



Discursos pronunciados en el Acto de Investidura
del Profesor
Mateo Valero Cortés
como
Doctor Honoris Causa por la Universidad de Murcia

Murcia
25 de abril de 2024

Universidad de Murcia
Servicio de Publicaciones, 2024

Depósito Legal: MU 412 – 2024

Imprime: Imprenta de la Universidad de Murcia



ÍNDICE

Manuel Eugenio Acacio Sánchez, <i>Laudatio in honorem del doctor Mateo Valero Cortés</i>	9
José Manuel García Carrasco, <i>Laudatio in honorem del doctor Mateo Valero Cortés</i>	17
Mateo Valero Cortés discurso de Investidura como Doctor Honoris Causa	23

Manuel Eugenio Acacio Sánchez

Laudatio in Honorem del profesor

Mateo Valero Cortés

Sr. Rector de la Universidad de Murcia,
Sr. Consejero de Medio Ambiente, Mar Menor, Universidades e Investigación de la
Región de Murcia,
Sra. Secretaria de Estado de Telecomunicaciones e Infraestructuras Digitales,
Autoridades académicas,
Compañeras y compañeros,
Amigas y amigos,

La Junta de la Facultad de Informática aprobó por unanimidad, en fecha de 31 de marzo de 2023, la propuesta impulsada por los tres departamentos de Informática para el nombramiento del profesor Mateo Valero Cortés como Doctor Honoris Causa por nuestra Universidad. Además, nos encomendó al profesor García Carrasco y a mí mismo el gran privilegio de poder actuar como padrinos en esta ceremonia, tanto por nuestra cercanía académica como personal con el profesor Valero.

Concurren en el candidato un serie de méritos de extraordinaria relevancia a nivel científico, de liderazgo y de difusión de la disciplina de la Informática en general, y de la Arquitectura de Computadores más en particular, por los que creemos muy adecuado el ingreso del profesor Valero en el claustro de honor de la Universidad de Murcia. A través del laudatio que promulgamos, trataremos de poner de manifiesto la relevancia definitiva de estos méritos.





Comenzando con el plano estrictamente científico, la labor investigadora del profesor Mateo Valero ha estado principalmente concentrada en el ámbito de la Arquitectura de Computadores y en su aplicación al diseño de Supercomputadores más eficientes y capaces.

Todos somos conscientes del ritmo vertiginoso al cual ha venido avanzando la industria de la Informática desde que a finales de los años 40 se pudiese en marcha ENIAC, el primer computador electrónico de propósito general, y que fue concebido originalmente para ser empleado en el cálculo de tablas balísticas por parte del Ejército de los Estados Unidos. Pues bien, desde entonces y hasta la fecha actual, la capacidad de cálculo de los ordenadores se ha ido acrecentando de manera frenética y sin precedentes teniendo en cuenta la evolución acaecida en otros ámbitos de la Ingeniería. De hecho, un teléfono inteligente contemporáneo, de los que podemos llevar cualquiera de nosotros en el bolsillo, tiene mayor capacidad de cómputo en su interior que un supercomputador de hace tan solo 20 años. Tal y como vienen ilustrando desde hace tiempo en una de sus obras de cabecera los divulgadores más reconocidos en la materia de la Arquitectura de Computadores, los profesores Patterson y Hennessy, si la industria del transporte hubiera seguido el ritmo de la industria informática, por ejemplo, hoy podríamos viajar de Nueva York a Londres en un segundo por poco más de 1 céntimo de euro.

Gracias a disponer de ordenadores cada más capaces se han podido ir abordando progresivamente retos científicos que unos años antes hubieran podido parecer ciencia ficción. Un ejemplo muy claro lo estamos viviendo en la actualidad con las increíbles aplicaciones de la inteligencia artificial generativa, con su exponente más conocido, ChatGPT, y que han sido posibles gracias a, entre otras cosas, la tremenda capacidad de cómputo de que disponemos hoy en día.

En esta carrera sin pausa hacia la construcción de ordenadores cada vez más rápidos se han aliado dos fuerzas impulsoras fundamentales. Por un lado, los avances muy importantes logrados en la tecnología empleada para la construcción de los ordenadores, que han posibilitado el disponer en un chip de elementos, los transistores, cada vez más eficientes y pequeños y, por lo tanto, en mayor número y más rápidos. Y por otro lado, la innovación que se ha ido realizando en el ámbito de la Arquitectura de Computadores a lo largo de los años, que ha permitido organizar los miles de millones de estos «ladrillos» que están disponibles a día de hoy en un solo chip, de forma óp-

tima para poder materializar las mejoras habilitadas por la tecnología en mayor capacidad computacional. Y es aquí, en este ámbito clave en el desarrollo de las Tecnologías de la Información y las Comunicaciones, el de la Arquitectura de Computadores, donde el profesor Valero ha venido desarrollando su labor científica durante más de 40 años y en el que ha tenido un rol protagonista internacionalmente reconocido.

De una manera muy resumida, podríamos decir que los esfuerzos de investigación desarrollados por el profesor Valero a lo largo de este tiempo han influenciado de forma decisiva el diseño y la ingeniería de partes importantes de los procesadores empleados actualmente en diversos ámbitos, que van desde la computación móvil (teléfonos inteligentes) hasta los supercomputadores más rápidos del mundo, pasando por los usados en computadores personales y videoconsolas. Estos elementos (los procesadores), como bien es sabido, constituyen el núcleo de cualquier sistema de cómputo ya que están encargados de la ejecución de los programas.

Pero permítanme la licencia de emplear algunos términos más técnicos para poder poner en valor con la rigurosidad debida la dimensión de la obra científica del candidato.

Así, las aportaciones realizadas por el profesor Valero en el campo de los procesadores vectoriales, estilo arquitectónico este que se encuentra presente en prácticamente todas las CPUs y aceleradores actuales (como las GPUs modernas), son universalmente reconocidas por investigadores de universidades y empresas como las mejores a nivel internacional. Los resultados de su investigación han inspirado varios diseños comerciales de empresas como Intel, IBM, NEC, Fujitsu o ARM (entre otras) y han permitido mejorar el procesamiento de las aplicaciones que manejan audio y vídeo, así como acelerar también el procedimiento de inferencia de las redes neuronales profundas.

Muy significativas, también, han sido las aportaciones realizadas por el candidato en el ámbito de los procesadores superescalares, paradigma este que domina el diseño de las CPUs actuales, y cuyos resultados han permitido solventar algunas de las limitaciones que iban siendo encontradas para poder avanzar en el número de instrucciones ejecutadas por unidad de tiempo.

Pero la labor del profesor Valero no ha estado concentrada únicamente en el diseño del hardware necesario para la ejecución de los programas (estos procesadores que mencionaba) sino que también se ha ocupado de las diversas capas software que aseguran un aprovechamiento óptimo del hardware





que hay por debajo. Así, el candidato ha realizado contribuciones muy destacadas en tecnologías de compilación, las cuales definen los programas encargados de traducir los programas escritos por los humanos a la representación que manejan los procesadores de los computadores. Así mismo, es preciso hacer referencia a las propuestas realizadas por el profesor Valero en los denominados sistemas en tiempo de ejecución, o lo que es lo mismo las capas de software que controlan la ejecución de los propios programas. Por último, destacar también el trabajo de investigación pionero desarrollado por el candidato a principios de los años 80 en la búsqueda de estrategias que permitiesen lograr una ejecución eficiente de diversas aplicaciones en un modelo arquitectónico muy de moda por aquel entonces, como era el de las denominadas arquitecturas sistólicas, y que poco después quedó relegado al ostracismo y encerrado en un cajón por su aparente falta de aplicación práctica. Curiosamente, y con la irrupción de las aplicaciones de inteligencia artificial basadas en el procesamiento de redes neuronales, estas arquitecturas, que presentan una serie de características que las hacen muy adecuadas para dar soporte a estas cargas de trabajo, han sido resucitadas y, ahora sí, han encontrado su razón de ser, siendo la base de los procesadores para redes neuronales profundas que están siendo construidos en la actualidad.

La obra científica del candidato la configuran más de 700 artículos de investigación que, en total, acumulan unas 18.000 citas, resultando en un índice h de 71 a fecha actual, lo cual es muy alto en el campo de la Ingeniería Informática. Además, el profesor Valero ha participado en la organización de más de 300 conferencias internacionales y ha impartido más de 800 conferencias invitadas, algunas de ellas en nuestra Facultad.

Fruto de sus importantes logros científicos son también los numerosos y prestigiosos reconocimientos, tanto a nivel nacional como internacional, que el profesor Valero ha recibido a lo largo de su trayectoria profesional. De hecho, el candidato fue seleccionado en noviembre de 2008 como uno de los 25 investigadores más influyentes en Tecnologías de la Información, tomando como referencia el periodo de tiempo comprendido entre los años 1983 y 2008. Repasar todos sus premios y distinciones de manera exhaustiva llevaría bastante más tiempo del que estimamos razonable para este laudatio. Así pues, concentraremos la atención en aquellos que consideramos los más destacados.

El profesor Mateo Valero ha sido galardonado con los tres reconocimientos más importantes a nivel mundial en el campo de la Arquitectura de Computadores: el premio Eckert-Mauchly, otorgado en el año 2007 por las sociedades

ACM e IEEE, «por su extraordinario liderazgo en construir un centro de investigación de Arquitectura de Computadores de talla mundial, por las contribuciones fundamentales en las áreas de computación vectorial y multi-hilo, y por ser pionero en nuevos enfoques básicos para el paralelismo a nivel de instrucción»; el premio Seymour Cray, otorgado en el año 2015 por IEEE, «en reconocimiento a sus contribuciones fundamentales en el campo de las arquitecturas vectoriales, procesadores superescalares, procesadores multihilo y procesadores Very Long Instruction Word» y el Premio Charles Babbage, otorgado en el año 2017 por IEEE, «por sus contribuciones a la computación paralela a través de un brillante trabajo técnico, tutoría de estudiantes de doctorado, y la construcción de un increíblemente productivo entorno de investigación europeo». Además del profesor Valero, solamente dos investigadores a nivel mundial han sido reconocidos con estas tres distinciones hasta la fecha. Se da la circunstancia, también, de que el candidato es, por el momento, el primer y único europeo en recibir el premio Seymour Cray. Mencionar, por último, los doctorados Honoris Causa que le han sido otorgados por cinco universidades nacionales y otras cinco universidades internacionales.

En definitiva, Sr. Rector, Sras. y Sres. miembros de la Comunidad Universitaria, espero haber sido capaz de transmitir la relevancia de los logros científicos del profesor Mateo Valero, que creemos sin duda que justifican sobradamente su idoneidad para pasar a formar parte del Claustro de la Universidad de Murcia como Doctor Honoris Causa.

Muchas gracias.



José Manuel García Carrasco

Laudatio in Honorem del profesor

Mateo Valero Cortés

Sr. Rector de la Universidad,
Sr. Consejero de Medio Ambiente, Mar Menor, Universidades e Investigación de la
Región de Murcia,
Sra. Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales,
Autoridades académicas,
Compañeras y compañeros,
Amigas y amigos,

El Doctorado Honoris Causa es una tradición secular fuertemente arraigada que revela la importancia académica que las universidades conceden a estas Distinciones. Se confiere con carácter honorífico, reconociendo los méritos excepcionales en favor de la ciencia, la cultura o el arte que los galardonados han acreditado.

Junto a los relevantes méritos científicos, que de forma brillante y breve ha expuesto el Prof. Acacio, permítanme ahora exponer, también de forma concisa, el perfil humano y la fructífera relación y colaboración que ha tenido con la Universidad de Murcia nuestro doctorando Honoris Causa.

Las palabras que mejor definen al profesor Mateo Valero son *trabajo, excelencia y generosidad*. Su inquietud y pasión por el trabajo bien hecho, por su atención a los detalles y por el progreso de su área de trabajo en nuestro país, han sido siempre actitudes contagiosas. Y junto a eso, su manera de ser, su bonhomía y campechanía que proyecta sobre sus actividades profesionales, y





que han constituido un foco de atracción para un número importante de jóvenes investigadores alrededor de todo el mundo. El que haya querido seguir su estela, entre los que me encuentro, ha disfrutado siempre de su tiempo, de su atención y de su amistad.

Conocí al Prof. Valero en el año 1992 en Valencia, en la reunión anual española de los miembros del área de “Arquitectura y Tecnología de Computadores”, que en aquel momento se llamaban “Jornadas de Paralelismo” y estaban en su tercera edición. Mateo había sido en 1983 uno de los primeros catedráticos de nuestra área (téngase en cuenta que la informática era una disciplina muy reciente), era uno de los impulsores de dichas Jornadas, y ya empezaba a destacar por los proyectos internacionales en que estaba involucrado, así como por sus publicaciones. Mi director de tesis doctoral, el Prof. José Duato, aprovechó este evento para presentarme a Mateo, y comentamos diversas cosas que podríamos hacer en temas docentes y de investigación.

Por desgracia, hasta el final de esa década no hubo mucho más contacto con el Prof. Valero. Pero en el año 1999 tuve el honor de poder organizar aquí en Murcia (más concretamente en La Manga del Mar Menor), dichas Jornadas de Paralelismo. Quiero dejar aquí constancia de mi agradecimiento tanto al Prof. Valero como al Prof. Francisco Tirado por su confianza en nuestro recién creado y joven Departamento para poder traernos aquí dicho congreso y, como se dice ahora, poner en el mapa a Murcia dentro del área de la Arquitectura de Computadores. Mateo estuvo muy atento al buen desarrollo del congreso, facilitándonos la invitación del Prof. Wen-mei Hwu de la universidad de Illinois para dar la charla magistral.

La década del año 2000 representa el afianzamiento y desarrollo de nuestra área en la Universidad de Murcia. Junto a la inestimable ayuda del Prof. Duato, que ya venía ayudándonos desde hacía más de una década, el Prof. Mateo Valero tuvo una contribución muy relevante (junto a otros muchos que nos ayudaron y que no puedo enumerar en este breve *laudatio*). Mateo estuvo viniendo durante varios cursos académicos a dar un curso de doctorado, acudió también a impartir diversas charlas y conferencias a alumnos y profesores de nuestro departamento, y siempre estuvo disponible para presidir tribunales de tesis doctorales y de plazas universitarias.

Gran parte de la culpa de esta intensificación en las relaciones con el Prof. Valero la tuvo el Prof. Juan Ignacio (Nacho) Navarro, quien siendo amigo personal mío era uno de los principales discípulos de Mateo. Su insistencia para

unir nuestros departamentos ha producido grandes frutos, uno de ellos el acto que hoy nos reúne aquí. La muerte prematura de Nacho Navarro nos priva de tenerlo entre nosotros, pero seguro que desde el Cielo está siguiendo este acto, disfrutando, y velando (como él siempre hacía) porque todo salga adecuadamente.

La red de excelencia europea HiPEAC (Red de Excelencia en Arquitectura de Altas Prestaciones y Compilación) comenzó en el año 2004 a iniciativa del Prof. Mateo Valero, el cual fue durante los primeros años de la red su Coordinador. Hoy en día, la red cuenta con más de 2000 miembros de todas las universidades europeas, con un buen grupo de miembros afiliados de universidades americanas y asiáticas, y el congreso anual es uno de los referentes mundiales de investigación en la arquitectura de computadores. Gracias al apoyo de Mateo, la Universidad de Murcia pudo formar parte de la red desde sus orígenes, estando presentes en la reunión inicial (*kick-off*) que tuvimos en la bonita ciudad de Niza. De hecho, gracias a la visión de Mateo de crear esta red de excelencia en aquel lejano año 2004, la Unión Europea pudo *reengancharse* en la carrera de los Supercomputadores, teniendo en la actualidad 3 ordenadores en la lista de los 10 primeros del TOP500 (por cierto, uno de esos 3 es el ordenador Marenostrom que tenemos en el Barcelona Supercomputing Center—Centro Nacional de Supercomputación, que dirige el propio Mateo, y esto como español es un orgullo decirlo). También, y gracias a los miembros de la red HiPEAC, la Unión Europea ha emanado recientemente una Ley Europea de Chips para acelerar la autonomía europea en la fabricación de chips ante la creciente demanda de semiconductores en sectores clave como la automoción, los electrodomésticos y la omnipresente Inteligencia Artificial. Desde hace un par de años, el Gobierno de España está apoyando este objetivo por medio del PERTE Chip, que tiene como meta reforzar las capacidades de diseño y producción de la industria de la microelectrónica y los semiconductores en España. Este proyecto estratégico está previsto que movilice una inversión pública de 12.250 millones de euros hasta 2027 y active a su vez un importante volumen de inversión privada, siendo gestionado por la secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales, cuya secretaria de estado, María González, tenemos el honor de tener hoy con nosotros.

En la década del 2010 tuvimos un menor contacto con el Prof. Valero. En parte, por su intensa dedicación al BSC y a sus múltiples proyectos internacionales de investigación. Mateo nunca ha sabido decir que no a ninguna pe-





tición de colaboración que le han hecho, y sus múltiples viajes a los cinco continentes dan prueba de ello. En parte también porque el grupo de la universidad de Murcia se había afianzado en la investigación, y ya podíamos desarrollar nuestros propios proyectos. Permítanme que, a modo de ejemplo, detalle algunos de los logros de nuestro grupo, como el haber conseguido un proyecto “Consolidator Grants” del Consejo Europeo de Investigación en el año 2018 a una propuesta de investigación liderada por el Prof. Alberto Ros, siendo en aquel momento la primera vez que se otorgaba una *Consolidator Grant* a un profesor de nuestra universidad. Como otros ejemplos de esta madurez investigadora, puedo decir con orgullo que en las últimas convocatorias de las becas Ramón y Cajal nuestro departamento ha sido siempre escogido por algunos de los beneficiarios de estas prestigiosas ayudas (incluso hace 3 años el número 1 de los receptores de la Ramón y Cajal fue un antiguo doctorando de nuestro departamento), o que de los más de 30 doctorandos que tenemos en la actualidad un 35% son extranjeros, o que tenemos antiguos doctorandos trabajando en empresas del sector tan conocidas como Intel, ARM, Meta (Facebook) o Huawei.

No he dicho nada que Mateo no conociera, pero he nombrado estos ejemplos aquí para dejar constancia como gracias a la valiosa ayuda y a los consejos del Prof. Valero (y de tantos otros como el Prof. Duato) hemos podido desarrollar también nosotros un buen grupo de investigación. ¡Mateo, a ti y a todos muchas gracias!

Finalmente, y para resaltar aún más la faceta humana de Mateo, me gustaría añadir que en 1998 fue nombrado Hijo Predilecto de su pueblo natal, Alfamén, y que en el año 2005 se le puso su nombre al único Instituto público de dicha localidad.

En definitiva, Sr. Rector, Sras. y Sres. miembros de la Comunidad Universitaria, espero haber sido capaz de transmitir el perfil humano y la implicación con nuestra universidad que ha tenido el candidato, y que, junto a los méritos científicos ya expuestos por el otro padrino, el Prof. Acacio, nos lleva a pensar que justifican adecuadamente su idoneidad para formar parte del Claustro de la Universidad de Murcia como Doctor Honoris Causa.

Muchas gracias.

Mateo Valero Cortés

SIN UNIVERSIDADES NO HAY CHIPS

Palabras pronunciadas por el profesor
D. Mateo Valero Cortés
con motivo de su investidura como
Doctor Honoris Causa por la
Universidad de Murcia

*Rector Magnífico
Consejero de Medio Ambiente, Mar Menor, Universidades e Investigación de la Región de Murcia,
Secretaria de Estado de Telecomunicaciones e Infraestructuras Digitales,
Autoridades académicas,
Compañeras y compañeros,
Amigas y amigos,
Señoras y Señores*

Querido Rector: quiero empezar agradeciendo esta distinción tan apreciada que me otorga vuestra Universidad. El doctorado Honoris Causa es la más alta distinción académica que una Universidad puede conceder, y he de decirles que lo recibo con gran alegría y humildad. Dado que es un reconocimiento, fundamentalmente, a la investigación que he desarrollado a lo largo de mi vida, quiero compartir este gran honor con mis estudiantes de doctorado y colaboradores. Sin su ayuda no habría sido posible. Si me permiten, y con mucha humildad, les diré que he tenido la suerte y el honor de haber sido el director/codirector de 58 tesis doctorales. Pero esos doctorandos directos, (les llamo hijos científicos) han tenido y siguen teniendo una gran descendencia. En la actualidad somos una familia de más 1000 personas doctores, con siete generaciones. Y cada semana, creo que nace un nuevo miembro en alguna parte del mundo. Entre todos, hemos creado una escuela española distribuida en el campo de la Arquitectura de Computadores, y la Universidad de Murcia tiene muy buenos maestros, miembros de esta escuela. Tenéis arquitectos de computadores al más alto nivel mundial. He de decirles, que considero a esta “gran familia científica”, como el legado con el que quisiera fuera recordado como científico.





Deseo empezar expresando mi agradecimiento a todas las personas de la Universidad de Murcia, que han hecho posible esta distinción. En primer lugar, a los profesores José Manuel García Carrasco y Manuel Acacio Sánchez. Los conozco desde el año 1999, cuando organizaron las Jornadas de Paralelismo en la Manga y desde entonces hemos realizado muchas actividades juntas y creo que enseguida, nos convertimos en amigos. Entre esas actividades, recuerdo haber sido invitado por ellos para impartir enseñanzas en los cursos de doctorado de los años 2000 y 2004. Durante esa década, vi como el departamento de Arquitectura de Computadores, de la mano del profesor José Manuel García Carrasco, se iba haciendo muy potente y aproveché para que me acompañaran a impartir docencia a profesores de gran renombre mundial como pueden ser Yale Patt (Universidad de Austin en Texas), Gurindar Sohi (universidad de Wisconsin en Madison), Mark Hill (Universidad de Wisconsin, Madison, ahora en Google), Wen-mei Hwu (Universidad de Illinois en Urbana Champaign, y ahora en Nvidia), Stefanos Kaxiras (Universidad de Uppsala, Suecia) y Per Stenström (Universidad de Gothenburg, Suecia). Los tres primeros son premios Eckert-Mauchly que es el máximo reconocimiento mundial en nuestro campo, en honor a los que diseñaron y construyeron el Eniac. Desde entonces, he venido a la Universidad no solo para impartir conferencias sino también para formar parte de los tribunales de tesis y de las plazas académicas

Antes de seguir adelante, les diré que mis primeros contactos con personas de la Comunidad Murciana se remontan a hace más de cincuenta años. Como ustedes saben, soy ingeniero de telecomunicación por la Universidad Politécnica de Madrid. Acabé la carrera en el año 1974. Pues bien, durante mis años de estudiante en Madrid, hice muchas amistades con colegas. La mayoría eran aragoneses como yo. Pero tuve el honor de compartir durante cuatro cursos de la carrera con Francisco González Mené. “Paco”, como le decíamos los amigos, era natural de Lorca. Alumno muy brillante. Nos peleábamos por ver quien sacaba mejores calificaciones. Dedicó su trayectoria profesional a varias actividades; pero la que más le cautivó, fue la de ser gran experto internacional en el tema de defensa y seguridad nacional. Falleció hace unos pocos años, y los muchos que le conocíamos, lo recordamos con mucho cariño. También en Madrid, en el último curso de la carrera, nos impartió algunas clases el entonces jovencísimo profesor Francisco Ros Perán. “Paco”, de nuevo, era también natural de Lorca. En el año 1974 marchaba al MIT (Massachusetts Institute of Technology) para realizar una de sus dos tesis doctorales. A su

vuelta, desarrolló una carrera profesional increíble. Entre otras actividades, trabajó en varias empresas como Telefónica Y Qualcomm. Fue secretario de Estado de Telecomunicaciones. Tengo el placer de verme con él de vez en cuando, para hablar de los temas de los chips, y siempre aprendo de sus ideas y de sus consejos. Paco es un crack a nivel técnico y humano.

Hace alrededor de 15 años, tuve la suerte de coincidir varias veces con María González Veracruz. María es nacida en Murcia y es bioquímica. La conocí cuando era la secretaria de Innovación y Nuevas Tecnologías del PSOE en época del presidente José Luis Rodríguez Zapatero. Mantuvo esa actividad desde el año 2008 al año que mantendría hasta 2012. Durante esa época, colaboré con ella y el fallecido Alfredo Pérez Rubalcaba, por entonces vicepresidente del gobierno, en labores de promulgar la investigación a nivel español. Con María nos hemos vuelto a reencontrar hace dos años ya que es la secretaria de estado de telecomunicaciones. Como ven, he conocido a dos secretarios de estado de telecomunicaciones nacidos en la comunidad murciana. María, es ahora, la responsable, entre otros temas del PERTE chip. Impulsado por el gobierno de Pedro Sánchez, con 12.400 millones de euros. María está ayudando a que el PERTE sea un éxito en general y yo le agradezco en cada momento, los consejos y ayudas que me da a mi para potenciar el diseño de chips basados en RISC V en España.

Hace más de diez años, tuve la suerte de conocer a Juan María Vázquez. De Juan Mari no voy a decirles nada que no conozcan. Es una persona admirable. Desde el primer momento, hubo química entre nosotros y me siento muy honrado con su amistad. Con él siendo secretario general de ciencia e innovación y presidente de CDTI, entró en operación el supercomputador Marenostrum 4. A través de Juan Mari, tuve la suerte de conocer a Teodoro García Egea. Teodoro, es también ingeniero de telecomunicación por la universidad politécnica de Cartagena. A nivel político, llegó a ser secretario general del partido Popular. Siempre le agradeceré su amistad así como que siempre apoyara al BSC para llegar a ser lo que es hoy en día. Y a través de Teodoro, conocí recientemente a Juan Luis Pedreño Molina. Ingeniero de telecomunicación también, y catedrático de la Universidad Politécnica de Cartagena. Ha desarrollado muchísimas actividades relacionadas con la carrera, entre ellas, Decano del Colegio de Ingenieros de Telecomunicación, presidente de la Asociación de Ingenieros de Telecomunicación de la Región de Murcia, consejero del Colegio Oficial de Ingenieros de Telecomunicaciones y vicepresidente de la Asociación nacional de Ingenieros de Telecomunicación





y Portavoz adjunto del Grupo Parlamentario Popular en la Comisión de Educación y Cultura.

Les diré también que, hace unos años, tuve la suerte de encontrarme con Darío Gil. Nacido en Murcia estudió los primeros cursos de su carrera en Madrid y marchó al MIT para acabarla. Ha desarrollado toda su carrera en los estados unidos, siendo en la actualidad el director de los laboratorios de investigación de IBM. Es el puesto más alto alcanzado por ningún español en IBM. Desde allí, siempre ha potenciado la colaboración entre IBM y España y en concreto, con el BSC por lo que le estoy muy agradecido. Quiero decirles que hace unos días estuvo en Madrid con el ministro José Luis Escrivá, y sus secretarios de estado Maite Lledó y María González, para firmar un acuerdo de colaboración entre IBM y España en el tema de la Inteligencia Artificial. Basados en ese acuerdo, el BSC e IBM desarrollaremos de manera conjunta modelos LLM para potenciar y proteger los cuatro idiomas del estado español, así como el diseño de chips altamente competitivos para ejecutar aplicaciones de la IA en la red. Todo ello será realizado con el objetivo en mente de que las administraciones y empresas españolas puedan beneficiarse de sus resultados y puedan crear miles de puestos de trabajo de alta calidad técnica y económica.

Debido a mi amistad con Juan María Vázquez, he realizado varias actividades en Murcia. Y estas actividades me llenaban el corazón con una inmensa alegría. Me refiero a las charlas que Fina García Lozano organiza cada año en Segura de Molina. Fina es una persona que no sé cuándo descansa. Trabajadora infatigable dedica la mayor parte de su tiempo a trabajar por los niños de la comunidad. Entre sus objetivos, el de “que ninguna niña y ningún niño sin alimentación adecuada cada día”. No vean ustedes la sensación tan agradable y única de verme dando charlas en una sala llena con centenares de personas jóvenes preuniversitarios con ganas de recibir información, formación y consejos que les sirvan para tomar decisiones importantes en sus vidas como es que carrera elegir. Increíble. Y en cada una de las charlas que he dado, mi mensaje al final a los jóvenes siempre era el mismo: “podéis ser en la vida, aquello que queráis ser, solo depende de vosotros”.

Y para acabar esta parte, he de decirles que me encuentro super contento de ser Académico de la Academia de Gastronomía de Murcia. Quiero agradecer a la Academia y en su nombre a su presidente Alberto Requena, al recibir tan alto honor. Demostraron una gran generosidad, al nombrarme Acadé-

mico. Tenemos un chat donde intento seguir, sin mucho éxito, todas sus actividades, entre las que se encuentran, multitud de artículos que escriben sobre el tema. Con mis compañeros académicos he pasado algunas jornadas inolvidables fundamentalmente en Murcia, pero también cuando vinieron a visitar el Marenostrom y degustamos una buena paella y buen pescado en el restaurante 7 puertas de Barcelona. También, en uno de mis últimos viajes a Murcia, tuve el honor de conocer y compartir momentos muy agradables con Raimundo González Frutos, creador y dueño del mítico rincón de Pepe y que falleció hace unas semanas. Gracias, compañeros de la Academia por vuestra enorme generosidad,

La iniciativa de los profesores José Manuel García Carrasco y Manuel Acacio Sánchez recibió enseguida el apoyo del profesor Jesús García Molina, catedrático del Área de Lenguajes y del decano de la Facultad de Informática, profesor Antonio Flores Gil. Poco después, la candidatura fue apoyada por los tres departamentos de Informática de la Facultad: el de Ingeniería y Tecnología de Computadores (su director por entonces era Lorenzo Fernández Maimó y los profesores del área Alberto Ros y Juan Luis Aragón la apoyaron vivamente), el de Informática y Sistemas (su director por entonces era Jesualdo Fernández Breis) y el de Ingeniería de la Información y las Comunicaciones (su director era Humberto Martínez Barberá).

Esto propició que la Junta de Facultad, en sesión de 31 de marzo de 2023, aprobase presentar la propuesta al Rectorado por unanimidad de todos los presentes. De hecho, he de agradecer a toda la comunidad de la Facultad de Informática, profesores, personal de administración y servicios y estudiantes, por la ayuda prestada para que este acto académico sea una realidad.

El Rector de la Universidad de Murcia, Prof. José Luján Alcaraz, acogió la propuesta con gran entusiasmo, una vez que esta salió de la Facultad de Informática. Los miembros de la Comisión de Distinciones de la Universidad de Murcia le dieron una excelente acogida a la candidatura, la cual se aprobó el día 2 de noviembre de 2023.

Finalmente, el Claustro de la Universidad de Murcia aprobó la propuesta de concesión del Honoris Causa al Prof. Mateo Valero en sesión ordinaria celebrada el 30 de noviembre de 2023.

No puedo estar más agradecido a todas las personas e instituciones que han hecho posible esta distinción que me honra y que llevaré siempre con-





migo con humildad, honor y alegría. A partir de ahora, el claustro de profesores tiene un nuevo profesor, que se pone a su completa disposición para colaborar con ustedes en cualquier tema que soliciten. Gracias a todos.

En las Universidades españolas por las que he recibido esta distinción, es parte del protocolo, que el futuro Doctor Honoris escriba un documento que sea utilizado como parte de la presentación pública. He tenido el honor de haberlos escrito en varias ocasiones. Y en ellos he tratado sobre temas tan diversos como la arquitectura de computadores de altas prestaciones, con especial énfasis en los supercomputadores, de la importancia enorme de la educación en la estructura de un país, de cómo motivar a los jóvenes y de la necesidad que los países tenemos de generar ideas que puedan ser aplicadas para el bien común. En el caso de la Universidad de Zaragoza, mi querida tierra natal, también me permití hablar sobre mis orígenes, a partir de mi nacimiento en un pueblo muy pequeño que se llama Alfamén, y que está en la provincia de Zaragoza, y sobre mi paso por el internado de los padres escolapios de Zaragoza durante 7 años y de mi primer curso de la carrera en dicha universidad.

Para esta ocasión, pensé que sería razonable no repetirles conceptos que ya he escrito en las ocasiones anteriores y usado en muchas de mis conferencias. Y que sería más razonable que les comentara la importancia estratégica que tienen los chips a nivel mundial. Les resumiré diciendo que el diseño y fabricación de los chips más avanzados, usados en la fabricación de los supercomputadores y en el diseño de las máquinas donde se ejecutan las aplicaciones de la Inteligencia Artificial, tales como el ChatGPT, son el tema geopolítico más importante a nivel mundial, si exceptuamos el horror de permitir que mueran miles de personas cada día por hambre y las guerras como en Ucrania y Gaza. Por otra parte, estamos en una Universidad y les comentaré que una vez más, nuestras instituciones tienen que recibir los recursos necesarios para formar ingenieras e ingenieros en estos temas estratégicos. De ahí que al documento que sigue le hemos titulado: “Sin Universidades, no hay chips”.

!!!Sin Universidades no hay chips!!!

Resumen

A lo largo de este documento se quiere comentar la gran importancia que tiene a nivel mundial el diseño y fabricación de chips de muy altas prestaciones; es decir de los más rápidos del mundo que se construyen usando los transistores más pequeños. Estos chips solo se fabrican en Taiwan, empresa TSMC, en Corea del Sur con la empresa Samsung y en Estados Unidos en la empresa Intel. Los chips más estratégicos hoy en día son los que se dedican a ejecutar los modelos grandes que necesita la Inteligencia Artificial. Hay una competición sin precedentes en este tema. En Europa no sabemos diseñar estos chips y dependemos totalmente del exterior. Europa se ha puesto las pilas para paliar un poco esta enorme desventaja tecnológica con la creación de la iniciativa Chip.act con 43 mil millones de euros. En España, el gobierno creó el programa Perte Chip con más de doce mil millones de euros, en total sincronía con la iniciativa europea. Muchas actividades se pueden hacer con ese dinero. Nuestra propuesta es que las dos más importantes, y las menos costosas a la vez, son la creación de talento en diseño de chips a través fundamentalmente de las Universidades, en colaboración con las empresas y, en segundo lugar, la ayuda a la creación y atracción de empresas que puedan diseñar estos chips que empiezan a dominar y dominarán el mundo en los próximos 25 años.

Introducción

Estamos viviendo un momento sin precedentes en la historia de la humanidad, marcado por una revolución tecnológica en la que los computadores, los datos y la Inteligencia Artificial, están cambiando nuestra forma de interaccionar con todo lo que nos rodea, causando un impacto directo en nuestro





día a día, a nivel laboral, en las empresas, en la economía, en el ocio, en los tratamientos de salud, en la movilidad, en la forma de investigar y en la de relacionarnos con nuestros seres queridos. Cada poco tiempo, nos comunican resultados derivados de los avances tecnológicos que nos sorprenden muchísimo hasta a las personas que no somos ajenas a la tecnología.

El transistor, sin duda, es la base de todo este cambio. Y nunca un dispositivo cambió tanto al mundo en menor tiempo. *“La culpa no fue del Chachachá, sino que fue del transistor”*.

Todo esta revolución tecnológica y social empezó con el invento del transistor en 1947 en los Laboratorios Bell de Estados Unidos. El transistor reemplazó, rápida y exitosamente, a las válvulas de vacío en la construcción de los computadores, e inició una frenética carrera en el diseño de todo tipo de chips tales como procesadores, aceleradores, memorias y redes de comunicaciones, entre otros. Y sobre ese hardware se ejecuta una ingente cantidad de software de forma que juntos, han cambiado la sociedad de una manera impensable.

Desde su invención, el proceso de fabricación de los transistores ha permitido reducir sus tamaños drásticamente. Como ejemplo, en el año 1971, se construyó el primer microprocesador de cuatro bits, el Intel I4004. Contenía alrededor de 2300 transistores planos y cuadrados cada uno de ellos, con 10 micras de arista. En la actualidad, los chips más avanzados contienen transistores de un tamaño de 3 nanómetros por lado; es decir, en 50 años, se ha conseguido que, en la superficie empleada para fabricar un transistor en el año 1971, se puedan integrar ahora, casi diez millones. Increíble, pero cierto!!!

“Los transistores son el ladrillo básico de todo este mundo tecnológico”. Sin transistores, no hay procesadores, sin procesadores no puede ejecutarse el software que influye en cualquiera de nuestras actividades humanas; sin software y hardware adecuados, y sin el acceso a grandes cantidades de datos, no existiría esa Inteligencia Artificial que no deja de sorprendernos cada día. *“Sin transistores, en definitiva, no hay paraíso”*.

“Los semiconductores son la base de nuestra sociedad”. Los procesadores son los cerebros de esta sociedad fuertemente interconectada. Procesan toda la información y controlan una intensa comunicación, entre humanos, entre humanos y máquinas y entre máquinas. Los hay de muy diferentes tipos en función del uso que se les va a dar. Podríamos clasificar los procesadores por el tamaño de los transistores utilizados y por su complejidad en el diseño; ambos

conceptos suelen estar muy relacionados. En un lado del espectro, están aquellos procesadores dedicados a tareas fáciles, como pueden ser las de controlar frigoríficos, o sensores sencillos y que contienen pocos transistores de tamaños relativamente grandes, de tamaños de varias decenas o centenas de nanómetros. En el otro extremo, están los más rápidos del mundo usados para fabricar supercomputadores o para ejecutar aplicaciones derivadas de la Inteligencia Artificial que precisan una elevada velocidad de cálculo. Utilizan los transistores más pequeños, que en la actualidad son de tres nanómetros. Hoy en día, tenemos, por ejemplo, chips que, en ocho centímetros cuadrados de Silicio proveniente de la arena de cuarzo de las playas y de betas geológicas, contienen más de cien mil millones de transistores de cinco nanómetros y que conmutan o cambian de estado a más de 2 GHz (más de dos mil millones de veces por segundo). De hecho, hace unos días fue anunciada la GPU denominada Blackwell que tiene dos chips conectados y contienen 208 mil millones de transistores, con una tecnología de 3 nanómetros. Entre ambos extremos, se encuentran los procesadores de prestaciones intermedias utilizados, por ejemplo, en el Internet de las Cosas (IoT), en los teléfonos inteligentes, tabletas, computadores personales, coches y servidores.

Como ejemplo concreto, los supercomputadores se construyen mediante la conexión a alta velocidad y baja latencia, de miles y millones de chips procesadores y aceleradores de las más altas prestaciones. Los procesadores tienen memorias asociadas a ellos de forma que la memoria global de los supercomputadores es muy grande y así pueden ejecutar, rápidamente, problemas que necesiten gran cantidad de datos y muchísimas operaciones a esos datos. Todos estos componentes están conectados por una red de interconexión muy rápida que les permite intercambiar datos a muy alta velocidad. La red de interconexión exhibe una latencia muy reducida y un ancho de banda muy grande de forma que los procesadores pueden trabajar conjuntamente en la ejecución de un mismo programa, y así, reducir drásticamente su tiempo de ejecución, comparado con el tiempo que necesitaría un sólo procesador. Estas capacidades de cálculo y sus enormes memorias asociadas están permitiendo a los supercomputadores el análisis masivo de datos para hacer frente a grandes retos de la sociedad, como son, la lucha contra el cambio climático, la búsqueda de nuevas fuentes de energía, de nuevos materiales y de nuevos fármacos contra el cáncer, o la explosión reciente de la Inteligencia Artificial.





Los supercomputadores como el MareNostrum 5, que inauguramos el 21 de diciembre de 2023, en el Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), son importantes instrumentos que tienen la ciencia y la ingeniería para construir un mundo mejor. Son máquinas que permiten a los científicos e ingenieros, soñar, resolviendo problemas que hasta hace poco parecían imposible.

La evolución tecnológica se basa en la constante miniaturización de los transistores, tal como predijo hace 50 años Gordon Moore (ley de Moore, cofundador de Intel Corporation) y que han permitido construir supercomputadores como el Frontier, que sobrepasó, en el año 2022, la velocidad de 10^{18} operaciones aritméticas por segundo (Exaflops) sobre números reales codificados con 64 bits, y más de 50 Exaflops cuando ejecuta aplicaciones que provienen de la Inteligencia Artificial, que utilizan el formato bf16 para los datos.. Comparados con los computadores de antes de la Segunda Guerra Mundial que construía Konrad Zuse en Alemania, los supercomputadores actuales son 10^{18} veces más rápidos. Esa velocidad se obtiene juntado más de un millón de procesadores y haciendo que cada procesador individual actual vaya más de 10^{11} veces más rápido que los procesadores de Zuse. Para dar otro dato más reciente, diremos que una parte del MareNostrum 5 tiene aceleradores Hopper de Nvidia. Pues bien, cada uno de esos chips (de los que el MN5 tiene varios miles), están contruidos en un poco más de 8 centímetros cuadrados de Silicio, conteniendo 80 mil millones de procesadores, es tres veces más potente que la primera versión del Marenostrum 1, que instalamos en la capilla en el año 2004, ocupando una superficie de 120 metros cuadrados y una altura de más de dos metros, y que fue clasificado como el cuarto supercomputador más rápido del mundo. Volviendo al MareNostrum5, diremos que está formado por 4 computadores. El más rápido de ellos, está dedicado a la Inteligencia Artificial, se basa es 4480 aceleradores Hopper de los descritos anteriormente, y tiene unas velocidades de 163 Petaflops de velocidad máxima y 138 Petaflops cuando ejecuta la prueba del Linpack, que le permite ser clasificado como el número 8 más rápido del mundo. Ejecutando aplicaciones de la Inteligencia Artificial, su velocidad es de 4.48 Exaflops. El otro componente más grande está formado se basa en 6192 chips Sapphire Rapids de Intel, con 62 cores cada uno, que obtiene un a velocidad en la prueba Linpack de 40.10 Patflops, que lo clasificó como en número 19 más rápido del mundo. Es decir, el MareNostrum 5 tiene dos de sus cuatro componentes clasificados en las posiciones 8 y 19 a nivel mundial en noviembre de 2023

Poco tiempo ha pasado desde que el MareNostrum 5 ha sido inaugurado. De hecho, solo 4 meses. Y en ese tiempo, Nvidia ha anunciado, este mes de marzo de 2024, un nuevo chip que mejora tremendamente el Hopper. Se llama Brackwell. De hecho, es un doble chip conteniendo en total 208 mil millones de transistores. Uno de sus parámetros es que cada chip puede ejecutar 20 Petaflops en números de coma flotante de 8 bits (FP8) o el doble de operaciones cuando los números están codificados en FP4. La información existente indica que este chip puede ir cuatro veces más rápido que el Hopper en las fases de entrenamiento de la red neuronal y treinta veces más rápido haciendo inferencias. Al igual que el Hopper, este chip se conecta al chip denominado Grace, que contiene 32 procesadores superescalares potentes con filosofía ARM. A nivel de la memoria externa, cada chip contiene una memoria 3D, HBM3e de 192 Gigabytes con un ancho de banda de 8Terabytes por segundo. Pero este superchip no actúa solo. Nvidia permite conectar hasta 72 de ellos a través de la red de interconexión NVlink con un ancho de banda de 1.8 Terabytes por segundo, de forma que esos 72 chips pueden trabajar como si fuera uno de ellos, pero 72 veces más rápido. La máquina se puede configurar en una torre (rack), con los 72 Blackwell, obteniendo una velocidad de 720 Petaflops en números con formatos FP8, FP6 e INT8 (enteros de 8 bits), y 1.44 Exaflops con números codificados en FP4. La memoria asociada es de 30 Terabytes con un ancho de banda agregado de 600 Terabytes por segundo. Y el objetivo próximo es llegar a conectar, con el mismo ancho de banda de los 72, hasta 576 Blackwell; es decir, ocho veces más.

¿Quién manda en los chips de altas prestaciones?

Para que los chips sean una realidad, hay que diseñarlos y fabricarlos. Usando un símil con la construcción de edificios, diseñarlos es como hacer los planos para que sean utilizados posteriormente, como guía para su construcción; es ahí, en este proceso de diseño, donde se define su funcionalidad. En el caso de los chips, hay que diseñar su estructura interna para “guiar” su producción en las fábricas denominadas “foundries” (coloquialmente, yo las llamo “hornos”). El poseer la tecnología más avanzada en el diseño y fabricación de chips, es tal vez la cuestión geopolítica más importante a nivel mundial, después de dejar morir de hambre a miles de ciudadanos del mundo y de las terribles y cruentas guerras en lugares como Ucrania y Gaza. *“Los chips más complicados dominan el mundo digital y, posiblemente, son las máquinas más complejas que ha creado la humanidad”.*





Pero no debemos de olvidar, siguiendo el símil, que las casas, los rascacielos, se hacen para que puedan albergar personas, para que alguien los utilice. En el caso de los chips, lo importante es que haya software que se pueda ejecutar eficientemente en ellos; que haya personas que se sientan cómodas en las casas. Y la falta de software a ejecutar, es el freno más importante que hace muy difícil el diseñar procesadores con juegos de instrucciones diferentes a los de Intel, AMD, ARM o Nvidia.

Sólo hay dos empresas en el mundo que pueden, a la vez, diseñar y fabricar los chips de mayores prestaciones: Intel y Samsung. En cuanto a las foundries, los países que poseen las fábricas con tecnología puntera son Taiwán, con la empresa TSMC (Taiwan Semiconductor Manufacturing Company), y Corea del Sur con Samsung. En estos momentos, TSMC está fabricando chips con una tecnología de 3 nanómetros y espera utilizar/conseguir 2 nanómetros al final del 2024 o principios de 2025. Estamos ya cerca del final de la famosa “ley de Gordon Moore”, ya que el número de átomos de Silicio, necesarios para fabricar esos tamaños de transistor, es tan reducido que empiezan a aparecer fenómenos cuánticos que hacen imposible fabricarlos. Además, el coste de cada una de esas fábricas avanzadas es superior a los 20 mil millones de euros. Importante hemos de mencionar aquí que el negocio relacionado con los semiconductores es superior a los 700 mil millones de dólares por año, alrededor de la mitad del PIB español del año 2023, y que, según McKinsey, superará el billón europeo, (millón de millones), antes del 2030.

La situación más habitual a nivel mundial es que haya empresas que quieren producir los chips de cualquier tipo, especialmente los más complejos, pero que no poseen esas foundries/fábricas/hornos. O sea, carecen de centros de fabricación. Ejemplos concretos diseñando chips de muy altas prestaciones, son Apple y Nvidia. Actualmente, hay más de 500 empresas subcontratando la fabricación de sus chips a TSMC. Son empresas de diseño denominadas “fabless” (sin fábricas/foundries). Indudablemente, ellos diseñan sus propios chips, pero en colaboración con las “foundries”, ya que necesitan coordinar sus diseños con todo el proceso de fabricación de chips; es decir, han de usar el costosísimo software de diseño, EDA (Electronic Design Automation), que hace posible compatibilizar el tipo de chips que quiere el arquitecto de los computadores con el proceso de fabricación utilizado en las “foundries/hornos”.

Ante el dominio de las “foundries” asiáticas sobre la alta tecnología, USA ha puesto el grito en el cielo ya que no puede admitir que, aunque sea el país donde se hacen los diseños del 80% los chips más complejos, y donde se han

desarrollado casi todas las herramientas EDA de diseño que utilizan en todas partes del mundo, no tengan tecnología competitiva para fabricarlos y deba recurrir a las “foundries” asiáticas. Actualmente USA diseña el 70% de todos los chips, sin importar la complejidad, y fabrica solo el 12% de la producción mundial de chips, mientras que ese porcentaje de fabricación era de casi el 40% en el año 1990. En la actualidad, Taiwan fabrica el 65% de los chips producidos a nivel mundial. Y de ahí que el gobierno de USA haya decidido, para continuar con su claro liderazgo en el tema de los semiconductores, invertir 280 mil millones de dólares en su programa Chip Act. También ha utilizado toda su “*diplomacia científica y de cualquier otro tipo*” para conseguir que TSMC construya dos “foundries” en Arizona con tecnologías de 4 y 3 nanómetros, y que Samsung construya otra en Austin con tecnologías similares. Las dos plantas en Arizona tendrán unos costes de 12.000 y 40.000 millones de dólares respectivamente. Cuando ya tenga estos clones de la empresa TSMC en Arizona, la cuestión es si USA estará tan fuertemente interesada en proteger a Taiwan ante una posible invasión de la China continental. Se calcula que la de Samsung en Austin costará casi 20.000 millones de dólares. Paralelamente, Intel ha anunciado unas inversiones de 20.000 millones de dólares para una fábrica en Ohio.

En el momento en que todas esas fábricas entren en funcionamiento, USA habrá recuperado su hegemonía tecnológica. Además, USA está inyectando gran cantidad de dinero para que Intel pueda competir de nuevo con TSMC y Samsung, ya que dejó de hacerlo, hace unos pocos años, en la carrera por los 12 nanómetros. Ahora, Intel quiere ser la primera fábrica que tenga tecnología de tamaños de transistores de 1.8 nanómetros, la tecnología denominada I18. Por otro lado, USA ha hecho lo imposible para evitar que China tenga acceso a la tecnología básica utilizada en las “foundries” más avanzadas como son las máquinas de litografía avanzada fabricadas en la empresa ASML, en Holanda, así como en el acceso a la compra de los chips más punteros. De esta forma, China lleva un cierto retraso en el diseño de las “foundries” más avanzadas. Ello ha hecho que empresas como Huawei ya no tengan acceso a los procesadores más evolucionados, necesarios en los teléfonos de altas prestaciones, y como consecuencia, ha pasado de ser líder mundial en ventas a una posición marginal. Hace unos días, sin embargo, China sorprendió al mundo y más en concreto a Intel, volviendo atrás un acuerdo para que Intel dejara de fabricar una foundry en China, de tecnología mediana, y por lo tanto no crítica para los intereses estratégicos americanos, por valor de casi 5000 millones de dólares. Por otra parte, China ya ha alcanzado autonomía total para diseñar y





fabricar chips con tecnología de cinco nanómetros. Y se acercan, rápidamente, a los reducidos tamaños de transistores de los competidores.

A título de anécdota, recordaremos aquí, una vez más, que la falta de chips para los coches y otros dispositivos, acaecida al final de la pandemia, fue una falta de planificación de los fabricantes de chips para coches al dejar de pedirlos a las fábricas, ya que redujeron drásticamente el número de coches fabricados. Al acabar la pandemia y volver a solicitarlos, las fábricas de chips tenían multitud de pedidos para fabricar dispositivos tales como teléfonos inteligentes y computadores personales. No fue ningún problema tecnológico, si no de falta de previsión de los fabricantes de coches, y de otros muchos equipos como frigoríficos y lavadoras, entre otros. Sin embargo, esta escasez en el acceso a cierto tipo de procesadores afectó de manera muy negativa a todas las economías a nivel mundial. Solo en USA, el departamento de Comercio evaluó las pérdidas en su país en 250 mil millones de dólares. Por ello, existe una iniciativa de este departamento de comercio que dedicará 50 mil millones de dólares en cinco años, de los que 11 mil millones serán dedicados a Investigación y desarrollo y 39 mil millones a potenciar el desarrollo de chips para ser usados en los Estados Unidos.

Sin talento, no hay nada que hacer

“Sin talento no hay ni investigación, ni innovación”. A nivel mundial hay una carrera tremenda en el terreno de formar cuantas más personas mejor en los temas asociados a las carreras STEM (Science, Technology, Engineering and Mathematics). De hecho, hay previsiones indicando que para el año 2030, China contará con hasta el 60% a nivel mundial, India con hasta el 20%, y Rusia, USA y Europa con menos de un 5% cada uno. A España se le predice el 0.7% Curiosas las predicciones para Indonesia (4%) y Arabia Saudita (3%), En ese contexto, el talento, en este campo, es el bien más apreciado: es escaso y está muy buscado y valorado económicamente. Hay una competencia mundial por la formación, retención y atracción del talento en todos los temas relacionados con el diseño de chips y la Inteligencia Artificial. A nivel mundial, hay un enfoque renovado en la autonomía digital para la economía digital. Todas las regiones y países avanzados están invirtiendo fuertemente en las capacidades para diseñar y fabricar semiconductores para impulsar sus economías digitales. La base de estos esfuerzos es la mano de obra especializada necesaria para diseñar y fabricar

semiconductores. La educación y la capacitación son fundamentales para permitir una industria de semiconductores próspera. *“Esto requiere la colaboración entre la Universidad, la Industria y el Gobierno”*. Recordemos que esta colaboración entre las administraciones, las empresas y las universidades fue fundamental para la creación del valle del Silicio en California. Las universidades como Stanford y Berkeley fueron las que ayudaron a crear el conocido *“Silicon Valley”*. Comentemos un poco este punto.

La invención del transistor se realizó en la costa este en los laboratorios de la Bell, y sus tres inventores fueron William Shockley, John Bardeen y Walter Brattain. Unos de ellos, William Shockley, decidió mudarse a la costa este, cerca de Palo Alto, que era donde había nacido, para seguir desarrollando allí sus ideas y crear empresas. Lo primero que hizo fue contactar con la Universidad de Stanford y potenciar las enseñanzas en microelectrónica para formar buenos ingenieros que permitieran mejorar el diseño y la integración de los futuros transistores. Vio la oportunidad que brindaba el nuevo dispositivo y aprovechó la coyuntura de tener Stanford, Berkeley y Caltech allí mismo. A partir de ahí, aparecieron empresas como Fairchild, que fue la primera en establecerse en el Valle del Silicio, y le siguieron otras muchas como Intel. Sin lugar a duda, al excéntrico Shockley se le reconoce como la persona que llevó el Silicio al Valle del Silicio. USA ha continuado desde entonces siendo el líder mundial en el diseño de chips de muy altas prestaciones. En la actualidad, tal como hemos comentado anteriormente, más del 70 % de los chips se diseñan en USA, aunque los más avanzados se fabrican en Taiwan y Corea del Sur. Las Universidades y empresas norteamericanas, apoyadas por el gobierno, han creado un fortísimo ecosistema que domina el mundo económico asociado a los chips y sus múltiples aplicaciones. USA creó el ecosistema perfecto entre administraciones, universidades y empresas y por ello, desde hace años, sus ideas, diseños y productos dominan el mundo.

El control y dominio en el diseño de chips es el tema más importante en la geopolítica mundial

Estados Unidos es el país líder indiscutible global. Pero, como hemos comentado, ha visto en peligro esa supremacía y ha decidido incentivar todos los aspectos relacionados con los semiconductores. Quieren dedicar 280 mil millones de dólares en 10 años, dentro de la iniciativa Chip Act. Lo importante





a resaltar, relacionado con el título de este documento es que la inmensa mayoría de esta cifra enorme, casi 200 mil millones, serán dedicados a la investigación y a la formación de talento en las Universidades. Asociado a la Chip. Act, se ha iniciado un programa super potente, dedicado a todo lo relacionado con semiconductores, con el objetivo de generar más y más talento en las Universidades. En concreto, piensan que necesitan durante los cinco años próximos, más de 50.000 nuevos ingenieros. De hecho, calcularon que, durante el año 2022, se generaron 20.000 puestos de trabajo en estas tecnologías. Con ese objetivo en mente, van a dedicar 174 mil millones de dólares en cinco años, para distribuirlos entre varias agencias que potencien la formación de talento y la investigación. De ellos, 80 mil millones han sido asignados a la NSF (“National Science Foundation”).

Además de esta gran inyección de millones desde la administración, existen grandes inversiones desde las empresas hacia las universidades y centros de investigación. Es impresionante ver, por ejemplo, como la existencia de las fábricas tecnológicamente más avanzadas, como la de Intel en Ohio, genera puestos de trabajo en el entorno próximo. Y por ello, los fabricantes son conscientes de que estos profesionales han de venir de la Universidad. En concreto, Intel está financiando con 50 millones de dólares muchas actividades universitarias, entre ellas, la cofinanciación del CASFRU (Center for Advanced Semiconductor Fabrication Research and Education). Programas similares se han establecido en multitud de Universidades como Purdue, West Lafayette, Florida o Illinois Urbana-Champaign, entre otras. La futura fábrica de TSMC en Arizona, ha empujado nuevos ímpetus en la formación de ingenieros en la ASU (Arizona State University), tal como la hizo, hace unos años, la instalación de una fábrica de Motorola. Es tan crítica la formación de profesionales en estos temas en Estados Unidos, que hasta han incorporado estudios de dos años en sus “Colleges”, para intentar cubrir la demanda. Ante su futura fábrica en Austin, Samsung también está potenciando la formación de talento; entre otras acciones, pagan la matrícula y libros a los mejores estudiantes de las Universidad de Texas. Hay varios ejemplos similares con otras empresas y universidades.

Por otra parte, Corea del Sur tiene en la actualidad un programa completísimo, uno de cuyos objetivos es capacitar a una fuerza laboral de 150.000 expertos para 2030. Empresas como Samsung han anunciado que necesitará con-

tratar a 80.000 personas conforme la demanda para chips de Inteligencia Artificial aumente. Corea del Sur, con sus empresas Samsung y Hynix, fabrica casi el 60% de las memorias, a nivel mundial, seguidas de la norteamericana Micron.

Corea del Sur, es un país que hace años apostó por la tecnología y lo sigue haciendo. El plan del gobierno, denominado “K-semiconductor strategy”), dedicará más de 450.000 millones de dólares hasta el año 2030, con el objetivo claro de crear la mejor cadena mundial de suministro de semiconductores. Hace años concentró su producción en las memorias. Con este nuevo plan, continuará en ello, pero, además, añade el reto de construir los mejores chips aceleradores para aplicaciones de la Inteligencia Artificial y técnicas para diseñar memorias donde se hagan cálculos también. En el año 2030, quiere que sus empresas de semiconductores exporten más de 200.000 millones de dólares.

Los cuatro actores básicos en los que se basa la iniciativa de Corea del Sur son: Gobierno, Universidades, compañías grandes y compañías pequeñas y recién creadas. Es esencial para ellos, la coordinación y colaboración entre los cuatro actores. El actor fundamental en esta iniciativa es la Universidad. Este ambicioso proyecto cuenta con el apoyo de 7 universidades con departamentos patrocinados por la industria y el gobierno para educar a estudiantes en el dominio del diseño de chips. Empresas como Samsung y Hynix empezaron hace años a financiar con grandes cantidades de dinero, la investigación en las Universidades, como las de Sungkyunk y Yonsey. Dan muchas becas a los estudiantes y les garantizan buenos puestos de trabajo al acabar sus estudios. Adicionalmente, el gobierno, en colaboración con las empresas anteriores, está invirtiendo 300 millones de dólares para la formación de los mejores estudiantes en sus Universidades. Importantes son, también, el instituto KAIST (Korea Advance Institute of Science and Technology) dedicado a las tecnologías de semiconductores en general, el AISS (Artificial Intelligence Semiconductor System) dedicado a la Inteligencia Artificial, y el PIM (Processor In Memory) dedicado a un tema de mucho futuro como es hacer que muchos de los cálculos se realicen en la propia memoria o cerca de ella, ahorrando cantidades enormes de energía al reducir el movimiento de datos entre las diferentes partes de los procesadores y sistemas computacionales. Al final del plan, el objetivo del gobierno es haber creado un entorno científico, económico y regulatorio, que promueva una colaboración real entre los cuatro actores descritos.





Corea del Sur era un país prácticamente agrícola, con una renta per cápita de menos de 1.000 dólares, en el año 1960. En la actualidad, su renta es de 38.000 dólares; es decir, en 60 años lo ha multiplicado por casi 40. Es conocido que sus empresas de semiconductores contribuyen con casi el 20% a su producto interior bruto. Ghana, otro país agrícola tanto en 1960 como ahora, tenía una renta per cápita de 1.000 euros en 1960, y de un poco más de 2.000 en la actualidad. Ponemos este ejemplo, como uno más en los que se demuestra que una buena decisión de los gobiernos puede cambiar la prosperidad de un país.

Junto con Corea del Sur, Taiwan, es tal vez, el país en el mundo que más cambios ha experimentado desde que sus políticos decidieron, hace no muchos años, potenciar el desarrollo de la tecnología de semiconductores. Como indicador económico, diremos que, en el año 1960, tenía una renta per cápita de 163 dólares y en el año 2022 de más de 30.000. Ha multiplicado ese indicador casi por 200. En el mismo periodo de tiempo, España ha pasado de tener una renta de 3.000 a casi 30.000 dólares. Lógicamente, las empresas de semiconductores de Taiwan son muy importantes para el país. Como ejemplo muy reciente, la taiwaniana TSMC inauguró el pasado 28 de julio un centro nuevo en Hsinchu, Taiwan, para potenciar la investigación y desarrollo de chips. Contará con 7.000 personas y una superficie construida de más de 300.000 metros cuadrados. Asociado a ese centro están las colaboraciones que la empresa fomenta con las mejores universidades de todo el mundo, ofreciéndoles acceso gratis a sus herramientas de diseño de chips y para hacer “tapeouts” (fabricar circuitos) con tecnologías muy avanzadas, de entre 7 y 12 nanómetros. Los precios para llevar a cabo esos desarrollos son muy reducidos. En Europa la conexión se hace a través de IMEC y Europractice. Pero no hay “cenas gratis”. Entre sus programas para atraer talento está la iniciativa de invitar a los mejores postdocs del mundo para que pasen, bien pagados, unos meses en sus instalaciones.

Por otra parte, China se dio cuenta muy pronto de la importancia de la arquitectura de procesadores RISC V, especialmente porque a pesar de su poderío económico e intelectual, había intentado desarrollar procesadores con un juego de instrucciones “made in China” y no lo consiguió. Es muy difícil, si no imposible, hasta para países como China, ir solos en este gran resto de diseñar procesadores que interpreten un juego de instrucciones propio y, sobre todo, desarrollar software que se ejecute en ellos. Como ejemplo concreto de que tener la última tecnología, en cuanto al menor tamaño en los transistores de los chips, es fundamental para competir, diremos que, en la última lista de los supercomputadores más rápidos del mundo, establecida en noviembre

del 2023 en Denver, el supercomputador número 11 de la lista es chino. Se llama Sunway y está construido con procesadores chinos. Su velocidad ejecutando el programa HPL que los clasifica en la lista es de 93 Petaflops. Diremos que el segundo de la lista, el norteamericano Aurora tiene una velocidad de 585 Petaflops (6 veces más) conteniendo menos de 5 millones de procesadores frente a los más de 10 millones del Sunway que son el doble.

En el año 2018, los chinos establecieron un serio programa hasta el 2030. Se denomina CRVA (China RISC-V Alliance), y tiene el objetivo de que, en el año 2030, China cuente con un entorno abierto de hardware y software, OSCE (Open Source China Ecosystem) El plan tiene tres partes que se van cumpliendo: la primera, durante los 5 años primeros el objetivo era diseñar chips con filosofía RISC-V y SoCs (Systems on a Chip). En este periodo se han desarrollado varios chips, algunos de ellos de altas prestaciones, como el de la empresa SOPHGO, que contine 64 cores superescalares con reloj de una frecuencia de 2 GHz. En paralelo, se diseña software de base como compiladores y aplicaciones enteras. En el segundo objetivo, que durará alrededor de 8 años, la idea es desarrollar herramientas software para ayuda a los diseños de chips, EDA (Electronic Design Automation). Estas herramientas son costosísimas y tener acceso a ellas gratis permitiría un avance muy importante. En el tercer objetivo, que durará hasta el año 2030, se pretende usar estas herramientas EDA para implementar procesadores de manera semiautomática, y otros dispositivos, de manera casi automática. El programa pretende que, en 2025, China tenga procesadores superescalares tan rápidos como los de ARM, Intel o AMD. Hasta hoy, más de 4000 estudiantes en China han participado en el programa OSOC. El número de estudiantes y de empresas, que esos estudiantes crean está creciendo de manera exponencial.

En la India, donde tuvimos la suerte de impartir unas conferencias en Bangalore en abril del 2023, nos informaron de los programas de sus universidades e institutos de investigación para formar miles de expertos cada año. Detectamos la energía que había en las discusiones técnicas relativas al diseño de los chips del futuro basados en el RISC V. India ha pasado de exportar talento, mayoritariamente a USA y Norte de Europa, a que las empresas multinacionales vayan a establecer sedes allí, tras una estructuración estratégica formativa como las iniciativas “Digital India” y “Made in india”. Las empresas van a donde se genera el talento. Las universidades forman ingenieros, y ese talento crea muchas empresas locales y atrae mucho dinero de las empresas extranjeras, que invierten allí para crear sus sedes.





Mirando hacia el futuro a nivel mundial

Tal como hemos comentado anteriormente, el tema del diseño de chips de altas prestaciones y de las fábricas donde se fabrican, es el tema más importante, geopolíticamente hablando, después de la gran hipocresía mundial que permite mueran miles de personas cada día, y de guerras tales como las de Ucrania y Gaza. Y este tema se ha complicado durante los últimos meses, coincidiendo con la aparición del ChatGPT y de otros modelos fundacionales. Estos modelos necesitan usar centenares y miles de los chips más potentes del mundo, junto con gran consumo de energía y la existencia de enormes cantidades de datos. La empresa Nvidia, está haciendo su agosto. Domina el 80% de los chips de altas prestaciones dedicados al entrenamiento y uso de estos grandes modelos. El precio de lista de cada una de sus potentes GPUs, denominadas Hopper las últimas, es de 30.000 dólares cuando el coste total de investigación y producción no llega a los 500 dólares.

ChatGPT utilizó, para entrenar la red, 10.000 chips Ampere de Nvidia, equivalentes en potencia a 3000 de los 4500 chips Hopper, que siguió al Ampere, y que coincide con los 4480 que contiene el MareNostrum 5, o a 700 de los recién anunciados Blackwell, que sigue al Hopper. Para fabricar una máquina, que permitiera obtener el 10% de la capacidad de cálculo de Google, se necesitaría una inversión inicial de 48 mil millones de dólares, con mil millones adicionales anuales para mantener las peticiones de los usuarios

Nvidia vendió medio millón de GPUs tipo Hopper (las más rápidas en la actualidad), en el tercer trimestre de 2023 a Meta (Facebook). Nvidia ganó más de 14 mil millones de dólares en el tercer trimestre del año 2023, y sus pedidos son tan grandes que acumulan un tiempo de servicio de entre 36 y 52 semanas. Nvidia vendió a Microsoft y Meta, en su conjunto, más de 150 mil de Hoppers y más de 50 mil a cada uno de Google, Amazon, Oracle y Tencent en el cuarto trimestre del 2024. En total, en ese cuatrimestre, Nvidia vendió más de 500 mil GPUs del tipo Hopper. Increíble, pero cierto.

ChatGPT es un gran modelo con muchos parámetros de entrada. Pero hay propuestas para hacer modelos mucho más grandes. Tengamos en cuenta que entrenar un modelo de un billón europeo de parámetros (un trillón de parámetros para los americanos), puede necesitar más de 10 elevado a 26 operaciones con números en coma flotante de 16 bits (formato bf16, propuesto por Google). Para decirlo de otra manera, necesitaría 204 días enteros del supercomputador

Aurora de Argonne, que es desde noviembre del 2023, el segundo supercomputador más rápido del mundo. Si se tuviera que alquilar ese servicio a Amazon, Google o Microsoft, el coste sería de 1000 millones de euros aproximadamente.

Toda esa supremacía de Nvidia y la necesidad de las grandes corporaciones como Google, Amazon, Meta, de utilizar la Inteligencia Artificial en sus servidores hace que éstas se sientan tentadas a fabricar sus propios chips aceleradores de Inteligencia Artificial para el entrenamiento de las grandes redes neuronales y para su uso (Amazon), así como en chips para hacer transcodificación de videos (Google) o para procesadores de video (Meta). Esta necesidad es triple: necesitan garantizar el acceso a los chips de altas prestaciones, necesitan pagar menos por ellos y necesitan diseñar chips más orientados a sus aplicaciones ya que los de Nvidia son, por el momento de bastante uso general, por lo que gastan más energía y son más lentos que los alternativos orientados a las necesidades de estas empresas. Nvidia, es muy consciente de esta situación, y ya ha empezado a diseñar chips en función de los grandes usuarios que se los compran. El mercado previsto para el año 2027 en servidores es de 200 mil millones de dólares y la tarta es muy apetitosa.

Como ejemplo muy reciente, diremos que Softbank está planteando dedicar 100 mil millones de dólares para crear una nueva empresa que lleve a cabo el proyecto Izanagi, con el objetivo de competir directamente con Nvidia. Recordemos aquí que Softbank es la propietaria de ARM, Ya pregunta inmediata que se nos plantea es si en el diseño del procesador se utilizará ARM o RISC V. En la aventura no solo participará Softbank con 30 mil millones de los 41 mil millones que tiene en la caja, sino que, además, participarán compañías fuertes del Medio Este. Y decir que esta iniciativa de Softbank es independiente de la que lleva con Sam Altman, fundador de OpenAI, que vamos a describir a continuación.

Y por si teníamos poco con las GAFAM (Google, Amazon, Facebook (Meta), Apple y Microsoft, ha entrado en el juego la empresa OpenAI creadora del ChatGPT. Ya hemos dicho la enorme capacidad de cálculo que necesitó para entrenar el modelo y para usarlo posteriormente. Y estos modelos se van a complementar/completar, con muchos más datos (textos, audios, videos, señales de sensores, resultados de las simulaciones de los supercomputadores). Y van a entrenarse modelos en todos los centros de supercomputación y en otros muchos laboratorios y empresas. La necesidad de cálculo va a crecer muchísimo. Además, estarán los servidores de las empresas y la enorme cantidad de puntos donde se hará “Edge computing” que requerirá un enorme





número de procesadores especializados en diferentes aplicaciones de la Inteligencia Artificial. En Europa se van a financiar las denominadas: “AI factories”, que van a ser centros con aplicaciones específicas de la Inteligencia Artificial. Poseen datos, necesitan computadores potentes para que, utilizando las técnicas de la Inteligencia Artificial, puedan obtener resultados competitivos en los múltiples sectores a los que se dedicarán. Como ejemplo concreto de esta iniciativa, Microsoft y OpenAI han llegado a un acuerdo, anunciado en marzo del 2024, para construir un centro de datos, que será operativo en el año 2028, y que costará cien mil millones de dólares. La máquina se llamará Stargate. Si solo contáramos el precio del hardware, estamos ante una máquina que podría ser entre cincuenta y cien mil veces más rápida que la torre que hemos descrito anteriormente con los 72 Blackwell, que superarían ampliamente en unos centenares de veces la velocidad del Zettaflops (10^{21}) acercándonos al Yottaflops (10^{24}).

El desarrollo de nuevas redes neuronales es una necesidad en los centros de supercomputación. Por ejemplo, los centros de supercomputación de Argonne, Riken y el BSC, junto con otros muchos centros, vamos a entrenar y usar modelos con más de un billón (europeo) de parámetros con el objetivo de utilizar estas redes en mejorar nuestras investigaciones en campos tales como la energía, el cambio climático y la medicina personalizada. Indudablemente, necesitamos reorientar el hardware que han de tener los futuros supercomputadoras de forma que ejecuten, de manera eficiente, las aplicaciones derivadas de la Inteligencia Artificial.

¿Y qué influencia tiene todo este tema de la explosión de la Inteligencia Artificial en el diseño y fabricación de chips? Muchísimas. Para empezar, recordemos que una foundry (fábrica) de chips de las últimas tecnologías con transistores de entre 2 y 3 nanómetros cuesta entre 20.000 y 30.000 millones de dólares. Solamente la máquina de litografía, fabricada por la empresa holandesa ASML, cuesta entre 300 y 400 millones de euros. Sam Altman, fundador y director de la empresa OpenAI, está liderando una cruzada con el objetivo de construir foundries, máquinas basadas en la energía atómica, y las mejores ideas en arquitectura de computadores y técnicas de diseño, para construir los millones y millones de chips avanzados orientados a ejecutar aplicaciones derivadas de la Inteligencia Artificial. No solamente es lo más aventajado en el diseño del software, sino que quiere competir en el mundo del hardware. Para ello, está intentado crear una bolsa de dinero de entre 3 y 8 billones (europeos) de dólares. Tengamos en cuenta que esa cantidad es entre

2 y 6 veces el PIB español. En comparación, la deuda corporativa de Estados Unidos en el año 2023 fue de 1.44 billones europeos y que la capitalización de Microsoft y Apple, de manera conjunta, es de 6 billones de dólares. Sólo Microsoft, ha aumentado su valor bursátil a 3 billones debido al efecto positivo de haber invertido fuertemente en OpenAI.

¿Y cómo podemos entender esta propuesta a nivel económico? Y sobre todo sabiendo que el mercado global de chips fue de 527 mil millones (527 billones americanos) de dólares en el año 2023 (con un coste de la maquinaria para producir esos chips de 100 mil millones de dólares) y que las previsiones para el año 2030, es de un billón europeo; es decir casi el doble.

La historia de la competición por el diseño de chips de muy altas velocidades para aplicaciones de la supercomputación y de la Inteligencia Artificial comenzó hace poco más de 10 años, cuando los chips de Nvidia empezaron a ganar mercado frente a los clásicos X86 de Intel. En la actualidad, el 80% del mercado de estos chips lo tiene cautivo Nvidia. Y todavía quiere más al diseñar chips más orientados a las necesidades particulares de usuarios como Amazon, Google, OpenAI o Meta. Les está mandando el mensaje de que Nvidia lo puede hacer mejor y más rápido en el tiempo y barato que ellos. Empresas como TSMC, Samsung e Intel, libran una competición feroz para lograr foundries con los transistores más pequeños. Las batallas y alianzas se multiplican cada día. Hynix es una empresa surcoreana que tiene más del 50% del mercado de las memorias de altas prestaciones. Recientemente, Hynix y TSMC, han hecho una alianza para desbancar a Samsung, en el desarrollo de las memorias HBM de la sexta generación. TSMC, Hynix y Nvidia han establecido otra gran alianza para conquistar el mercado global de la Inteligencia Artificial. Mientras estos gigantes ven la batalla desde muy arriba, naces pequeñas empresas como Tenstorrent, Cerebras, Sambanova, o la recién nacida, la española Openchip, spin-off del BSC, que intentar hacerse un huequecito. El futuro, es muy impresionante.

Europa y España

En Europa y España, la situación no es muy saludable. Diseñamos y fabricamos chips de no muy altas prestaciones, tales como los que se utilizan en los coches actuales. Sin embargo, ni diseñamos ni fabricamos chips de altas prestaciones. No existe una política coordinada, por ejemplo, ni para la creación de talento, ni para repartir las foundries entre los países.





Como hemos comentado al hablar de la creación de talento en otras partes del mundo, Europa (y España en particular) deben invertir en la educación y en la formación de la próxima generación de ingenieros e investigadores que puedan allanar el camino para un ecosistema de semiconductores vibrante. Hoy, Europa tiene algunos de los componentes, como Europractice, pero necesita invertir más y habilitar estas opciones para las pymes y la industria. Recientemente, se han hecho anuncios para expandir el apoyo europeo al desarrollo de semiconductores, pero los mecanismos no están bien definidos. Desafortunadamente, Europa no ha abordado el desafío de la fuerza laboral. Europa debe establecer programas similares para habilitar la fuerza laboral de semiconductores. La Unión Europea puede apoyar los esfuerzos de Francia (CEA), Alemania (HM, Fraunhofer), Italia (Universidad de Bolonia), España (BSC) y Suiza (ETHZ) para poner en marcha los programas.

En el tema de las foundries, las diferentes naciones miran este tema a nivel muy egoísta. Por ejemplo, Alemania quiere llevarse todas las foundries avanzadas que se puedan construir en Europa en un futuro próximo. Existen programas parciales de R+D a veces desconectados entre ellos. La política y la burocracia no dejan avanzar a Europa en estos temas. Vamos, un pequeño desastre. Un producto de todo ello lo constituye el caso de los supercomputadores. Europa posee supercomputadores muy rápidos, como el MareNostrum 5 que inauguramos el 21 de diciembre en el BSC. Pues bien, ni éste ni ningún supercomputador europeo tiene procesadores y aceleradores diseñados y/o construidos en Europa. Creemos que esta situación no debe continuar. Desde siempre, el BSC ha estado detrás de evangelizar y de motivar a Europa de que no podemos permitirnos esta dependencia tecnológica. Por ejemplo, los futuros coches autónomos basarán dicha autonomía, entre otras cosas, en la existencia de chips de muy alta velocidad. Europa ni los diseña ni los fabrica, por lo que vamos a ser totalmente vulnerables si no tenemos acceso a ellos. Y de igual forma, si no fabricamos los superchips que necesita la Inteligencia Artificial, así como su software asociado, no podremos influir en que la Inteligencia Artificial contemple los principios éticos europeos al máximo nivel. Sin la tecnología, Europa solo puede intentar ser *“el cuarto árbitro de partidos en los que sabe que no tiene equipo para jugar”*. Poco a poco, desde el BSC hemos influido a la Unión Europea de cambiar esta situación.

Reducir la dependencia tecnológica de Europa frente a las compañías norteamericanas y asiáticas, que dominan este mercado, es uno de los objetivos estratégicos de la Comisión Europea, que a través de la iniciativa Chip Act

europea y otras, busca una buena posición europea en el diseño y producción de todo tipo de chips. Europa dedicará 43.000 millones de euros a esta actividad que intentará potenciar la creación de fábricas con tecnologías avanzadas, y el diseño de los chips que necesitamos. En cuanto a casos concretos, TSMC va a instalar en Alemania, una fábrica en Dresden, orientada a la producción de chips para coches, entre otras aplicaciones. Es importante que empresas como Infineon, NXP y Bosch hayan invertido cada una de ellas un 10% en el proyecto y sobre todo, que hayan decidido fabricar los futuros chips para coches en ella. Por otra parte, se intenta que Intel construya otra fábrica con la última tecnología, también en Alemania, y que mejore las condiciones de la que tiene en Irlanda. Francia quiere tener una fábrica de “Global Foundries” con tecnología lo más avanzada posible

Los chips dominan el mundo, por lo que, si Europa confía en tecnología extranjera para nuestras infraestructuras más trascendentales, puede ser un problema desde un punto de vista geopolítico. Y no se trata sólo de tener tecnología europea sino también de garantizar que ésta sea competitiva en el mundo. Europa tiene que garantizar su soberanía tecnológica y, para ello, debe fabricar sus propios procesadores basados en hardware abierto. Esto, sin ninguna duda, va a mejorar nuestra competitividad y rendimiento económico, pero, además, nos va a permitir crear aplicaciones que se ajusten a los principios éticos, legales, socioeconómicos y culturales europeos.

Y ahora se dan unas condiciones únicas para intentar reducir esa dependencia tecnológica. Por una parte, estamos acercándonos al límite del tamaño mínimo de los transistores de Silicio, de forma que todo el mundo podrá tener acceso a la tecnología más avanzada. Ahora entramos en una cierta estabilización tecnológica, sólo que algunas empresas como Apple y Nvidia tienen prioridad, debido a sus grandes volúmenes de consumo. Por ejemplo, se sabe que Nvidia y Apple han obtenido de TSMC, la prioridad para que, en el primer año y medio, sean casi los únicos clientes de chips que usen los tamaños de 3 y 2 nanómetros. Este hecho, les da una gran ventaja competitiva para fabricar productos que otros no podrán tener. En segundo lugar, y más importante, es que existe una iniciativa mundial, denominada RISC V, que define un juego común de instrucciones del lenguaje máquina de los procesadores, de forma que nadie es propietario, y todo el mundo puede fabricar procesadores que usen estas instrucciones. Este hecho está cambiando el ecosistema en el diseño de procesadores. Hasta ahora, las empresas como Intel, Nvidia, ARM o IBM son las pro-





pietarias de los juegos de instrucciones del lenguaje máquina de sus procesadores, y esto, prácticamente, hacía imposible el desarrollo de procesadores con juegos de instrucciones nuevos, ya que no hay software desarrollado para ellos. Era misión imposible. Europa tuvo en los procesadores ARM, su única oportunidad hasta que dicha empresa fue adquirida por Softbank (Japan), <https://www.softbank.jp/en/corp/aboutus/message/>.

El RISC V es equivalente en hardware a lo que Linux fue en software. Linux democratizó el software haciendo que todo el mundo pudiera hacer programas que se puedan ejecutar en cualquier plataforma hardware. RISC V está haciendo lo mismo en hardware. Multitud de empresas y centros de investigación están desarrollando hardware y software basado en RISC V y Linux. Se abre una posibilidad jamás imaginada, que va a permitir la existencia de un mundo abierto en el que la colaboración sea lo más natural. En pocos años, la inmensa mayoría de los procesadores serán RISC V. Con estas dos condiciones anteriores, aparecerán innumerables nuevas ideas que podrán ser llevadas a la práctica abriendo un campo de innovación nunca imaginados. La idea RISC V permite diseñar procesadores orientados a las aplicaciones, y esta será una ventaja adicional. La tercera razón es que Europa de manera global, y algunos países como España a nivel particular, han decidido dedicar importantes recursos para intentar obtener esta autonomía tecnológica. La Unión Europea y España, estamos ante un tren que debemos tomar; será la última vez que pase.

¿Qué hacemos desde el BSC?

Desde el BSC convencimos a la Unión Europea para que financiara proyectos orientados a desarrollar chips de muy altas prestaciones, basados en RISC V, para ser utilizados en los futuros supercomputadores, coches autónomos, etc. El tema no fue fácil. Hubo negativas muy fuertes en el RIAG (Research and Innovation Advisory Group), que es un grupo de asesores para la Comisión Europea en temas de supercomputación, del que formo parte. El resto de los expertos decía que era demasiado pronto tomar esa decisión; que era muy arriesgado. El tiempo, desgraciadamente, ha dado la razón al BSC. Adicionalmente, a través de discusiones con otras partes de la Comisión Europea, los convencimos de que el RISC V era la única solución que tenía la Unión Europea para diseñar chips, desde los que se utilizan para el IoT (Internet de las Cosas) hasta los de los teléfonos, tabletas, ordenadores personales, servidores, coches

autónomos, Inteligencia Artificial y Supercomputadores. Y así lo aceptó la Comisión Europea en un Workshop que se realizó a finales del año 2019. Esa fue una gran noticia para intentar llegar a una autonomía europea en el diseño de chips: aunque falta muchísimo camino por recorrer. En la actualidad, existe la EuroHPC Join Undertaken que potencia el desarrollo de chips basados en RISC V para altas prestaciones. Por otro lado, existe la KDT-JU (Key Digital Technologies Join Undertaken) que potencia proyectos para desarrollar procesadores y aceleradores basados en el estándar RISC V.

Con mucho esfuerzo, empezamos tarde, en el año 2019 y con pocos recursos, el proyecto EPI (European Processor Initiative), de chips aceleradores, que ha producido los primeros buenos resultados. Le han seguido otros proyectos tales como el E-Processor, el MEEP o los pilotos con RISC V. Y otros van a continuar. Vamos avanzando en este tema, aunque la distancia que nos separa con los productos actuales de Nvidia, Intel o AMD es todavía enorme. El BSC es el centro de supercomputación referente en Europa para el diseño de chips para altas prestaciones, liderando hasta ahora todos los proyectos europeos sobre el tema. Queremos que el MareNostrum 6 que tendremos en 5 o 6 años, contenga ya, procesadores y aceleradores desarrollados en España.

España y los chips

En España, la situación es todavía mucho peor que la media europea. Somos consumidores netos de chips, apenas diseñamos chips sencillos y nuestras escasas fábricas, lo más avanzado está en el Centro Nacional de Microelectrónica, tienen unas tecnologías no actualizadas, y no competitivas (100 nanómetros). Dándose cuenta de la importancia de los chips y para actuar de manera coordinada con Europa, el gobierno de Pedro Sánchez, y en particular su vicepresidenta, Nadia Calviño, concibieron un plan que el exsecretario de Estado de Telecomunicaciones, Roberto Sánchez plasmó en la realidad, para dedicar 12.400 millones de euros, de los fondos de recuperación, en una iniciativa única en España en consonancia con la Europea del Chip Act. La iniciativa se denomina PERTE chip y está liderada, a nivel político, por María González, secretaria de Estado de Telecomunicaciones y a nivel técnico, por Jaume Martorell, persona con gran experiencia técnica en estos temas. Tengo el honor de estar en el comité de expertos, asesorando al gobierno en este tema.





¿Qué podemos hacer en España ante esta situación única que tenemos con la financiación a través del PERTE chip? El plan contempla 4 apartados, todos ellos alineados con los objetivos del Chip Act europeo, que van desde intentar construir foundries y/o líneas de diseño, como las anunciadas con Broadcom e IMEC, hasta la formación de profesionales, pasando por el apoyo a empresas nacionales, ya existentes, o creando nuevas, como Openchip, y a empresas internacionales que ya estén establecidas en España, como Cisco o Maxeler, o bien, que aprovechen la situación para hacerlo.

Vamos a describir, más en detalle, acciones concretas sobre dos temas que nos parecen fundamentales: *“la creación de talento y la oportunidad de potenciar la existencia en España de empresas de diseño de chip fabless”*.

Desde hace varios años, en nuestras charlas sobre chips, cuando hablamos de cómo potenciar el diseño de chips en España, tenemos un lema que vamos anunciando y defendiendo en cada reunión que tenemos: *“Sin Universidades, no hay chips”*. Las Universidades son la pieza fundamental en ese cambio que queremos para España. Las Universidades son las generadoras de talento, y es a ellas a las que hemos de recurrir. Hemos de dotarlas de los recursos necesarios para que formen a nuestras ingenieras e ingenieros. Y hemos de llevar también el plan de estudios a la Formación Profesional. Necesitamos, ya un plan de choque que nos permita formar al menos 1000 expertos por año. Tenemos una ocasión única para que en este nuevo entorno mundial de software (Linux) y hardware (RISC V) abiertos, nuestros jóvenes universitarios puedan colaborar desarrollando nuevas ideas y chips de manera conjunta. La formación de talento debería de ser, a nivel estratégico, el apartado más importante del Perte Chip español.

Con ese objetivo, creamos hace 3 años, una asociación denominada SOHA (Spanish Open Hardware Association), en la que colaboramos ahora 33 Universidades españolas (estando abierta a cualquier otra Universidad europea o latinoamericana), y 4 centros de investigación, para coordinar las enseñanzas relacionadas con el diseño de chips basados en RISC V que, como hemos dicho, es un entorno abierto, que todo el mundo puede utilizar sin pagar royalties a nadie. RISC V es en hardware lo que Linux fue en software. Con Linux y RISC V, tenemos un espacio único para que todo el mundo, en especial el académico,

podamos colaborar en el diseño de software y hardware abierto. La verdad, es que esto parece increíble y era muy difícil, si no imposible, de imaginar hace pocos años. Si España produce buenos profesionales en el diseño de chips basados en el RISC V, ellos crearán muchas empresas nuevas y empresas extranjeras se instalarán en España para producir chips, tal como está pasando en India. Los países más avanzados en el diseño y fabricación de chips, tal como hemos descrito anteriormente, tienen programas en sus universidades para formar talento. Y así ocurre en India, Corea del Sur, China o Taiwan, además de, por supuesto, Estados Unidos. Y aun teniendo programas especiales estos países necesitan más y más ingenieras e ingenieros en el diseño de chips.

Por tanto, necesitamos establecer planes de estudios coordinados por todas las universidades españolas, grados, másters y programas de doctorados centrados en “el diseño de chips”, donde profesores y alumnos de todas las universidades españolas compartan conocimientos. Desde el PERTE chip, hemos de dar a las Universidades españolas, los recursos necesarios. Y no hablamos de mucho dinero. Los 80 millones contemplados en el PERTE son un buen punto de partida para empezar a cambiar la situación española.

El segundo tema que consideramos de suma importancia es el facilitar que desde España se hagan diseños muy avanzados basados en RISC V. Estamos, de nuevo, ante una situación única. Creemos, tal como hemos comentado, que, dentro de 10 años, la mayoría de los procesadores serán RISC V (en la actualidad, ya son el 10%, aunque casi todos son de muy bajas prestaciones). Y todo está por hacer. Hemos de crear un clon mejorado de todos los productos que hoy producen Intel, IBM, ARM o Nvidia, con sus juegos de instrucciones propietarios en el mundo RISC V. Y los primeros que los consigan, van a dominar los mercados. Es un gran reto, pero las empresas lo harán. Y serán, tal vez, empresas que hoy no existen o de las que existen muy pocas. Como ejemplo, pensemos en el mundo de los servidores. Son miles de millones anuales que se dedican a comprar estos sistemas. Y que el mercado de chips para coches será en pocos años bastante superior al de los servidores. ¿Y qué decir de los sistemas orientados a ejecutar eficientemente los modelos más avanzados de la Inteligencia artificial? Los primeros que sean capaces de diseñar chips para este mercado generarán muchos puestos de trabajo y dinero. Es una carrera hoy abierta para sustituir toda la gama de procesadores existentes.

¿Qué podemos hacer en España? Intentar tener aquí algunas de las empresas que tengan como objetivo ser muy competitivas en algunos campos, como,





por ejemplo, los de los computadores de altas prestaciones, o en el de los servidores o coches autónomos del futuro. Ahora es el momento, y no tengo duda de que atraer a estas empresas debería ser, junto con la formación de buenos profesionales, el complemento ideal para retener y atraer talento externo. En la actualidad, el PERTE chip ya ha dado pasos en esa dirección con la potenciación de empresas españolas, con la creación de empresas nuevas y con la ayuda a empresas multinacionales que han creído en el talento español, y que han tomado decisiones de las que todos nosotros deberíamos estar contentos.

Nos gustaría que, en pocos años, tuviéramos en España, a través de nuestras excelentes Universidades, un ecosistema de generación de talento tan bueno como en cualquier otra parte del mundo, y un tejido industrial, asociado a lo anterior, que sirviera para generar miles y miles de puestos de trabajo de alta calidad. Habríamos puesto a España en el mapa de la tecnología que hoy domina y dominará el mundo. Estamos en el momento adecuado, y podemos hacerlo

Hay frases que hemos usado en este escrito que me gustaría resaltar, tales como: “la culpa no fue del chachachá, fue del transistor”, “sin transistores no hay chips”, “sin Inteligencia Artificial no hay paraíso”, y la que sirve de título al escrito: “Sin universidades, no hay chips”. Extendiendo esta última frase, creo que sin buena educación juvenil y sin buenas universidades no hay países avanzados. Hoy en día, tenemos una oportunidad única en España, para avanzar hacia un futuro digno para un país como el que somos y queremos ser. Sé que es un sueño, pero nos gustaría pensar que España puede ser uno de los países punteros en el diseño de los futuros chips basados en RISC V. Nos gustaría que España fuera el **“European Design Valley”**.

Querido Rector Magnífico, Autoridades Académicas, Distinguidos Doctores, Señoras y Señores, Amigas y Amigos,
Muchísimas gracias

Barcelona 4 de abril de 2024